DOCUMENT RESUME

ED 088 947 TM 003 511

AUTHOR Brown, Stephen W.
TITLE The National Teacher Examination as an Instrument of Prediction of Teacher Success.
PUB DATE Apr 74
NOTE 14p.; Paper presented at the American Educational Research Association annual meeting (Chicago, Illinois, April 15-19, 1974)

EDRS PRICE MF-$0.75 HC-$1.50
DESCRIPTORS *Effective Teaching; Elementary School Teachers; *Evaluation Criteria; *Predictive Ability (Testing); Principals; Secondary School Teachers; *Teacher Evaluation; Teacher Rating; Teaching Skills; *Test Validity
IDENTIFIERS *National Teacher Examinations

ABSTRACT

The intent of this study was to determine if any relationship existed between teacher effectiveness, based on selected criteria generally accepted by educators to dencte effectiveness, and National Teacher Examination, Teaching Area Examinations (TAE) and Weighter Common Examinatien Total (WCET) scores. Teacher effectiveness was determined by the building principal's evaluation. Analysis of the data was accomplished by determining relationship through the correlation technique. It was found that a significant relationship (.05 level) does exist between the Teacher Area Examination of the NTE and secondary principal's evaluation. Hcwever, no significant relationship was found between elementary principal's evaluation and any of the NTE scores. (Author)

ED 088947

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

In this day and age it is particularly important for achievement and proficiency in any field to be recognized and capitalized upon. This is no less true, perhaps more true, in teaching, where the tremendous pressures of accountability are being applied. Efforts to find new ways of identifying competences are certainly in order. External examinations as predictors of competence, if used properly, are among imaginative approaches that have been adopted increasingly over the past decade.

The National Teacher Examinations have been used to assess the knowledge of teachers and prospective teachers since 1940 when the examinations were first administered by the American Council on Education. In addition to using the examinations to measure teacher knowledge they, in many instances, have been used by school systems in an attempt to measure present teacher effectiveness, e.g., the 1971 report by the Alabama Education Study Commission that "indisputable evidence" (based on NTE results) provides information that ". . . many of Alabama's teacher college graduates are alarmingly deficient for a profession in which proficiency is absolutely vital (1971)." The utility and accuracy of ratings of teacher performance have long been attacked and questioned, and this is especially true of principal ratings. The National Teacher Examinations are essentially achievement tests. How good are they as predictors of future performance? Cromer, (1961, pp. 155-158) in his article on testing programs, states:

> Too often we fail to realize that a test score is best interpreted as a good estimate of the <u>general</u>

> level of performance, and that it will vary from test to test and from time to time. The assessment of human traits and abilities is not at the same level of accuracy as that found in a physics laboratory. It probably is closer to the level of accuracy found in the predictions of weather, in which temperature predictions are generally within a few degrees of actual temperatures, but in which differences of ten or more degrees are common enough to be remembered vividly by critics.

This author goes on to say:

> There is considerable overlap between standardized tests of achievement and standardized intelligence tests. One of the important differences between the two is the way the results are used. When analyzing achievement test scores one is generally looking forward to the future, predicting performance.

It is well to keep in mind the fact that intelligence is inferred from achievement. We have no direct measures of intelligence completely divorced from achievement.

With the above distinctions in mind, what about the National Teacher Examinations? These are achievement tests, but they are also used as the basis for predicting future achievement or performance. This circumstance accounts for some of the confusion and ambivalence one encounters concerning these examinations. Is the situation contradictory? Probably not. Teachers must have considerable education and preparation in order to perform successfully as teachers. Achievement is expected in a variety of ways at adequate levels of competence. Such achievement is not possible lacking sufficient basic intelligence but alone, without preparation and study and achievement, does not qualify one for the job to be done. Not just any intelligent person is by virtue of his intelligence qualified to be a physician, lawyer, architect,

teacher or what have you, without study and achievement related to the expectations of the particular profession. It follows that tests of achievement in areas related to the future work of a teacher provide useful information concerning the general level of performance that may be expected in those areas in the future. If this is not so, then we all live in a fool's paradise.

One of the strongest supporters on the NTE (Seashore, 1965) indicates the tests measure knowledge of subjects, professional understanding, and mental abilities which experts agree are a requisite for good teaching.

Others (Boozer, 1968 and Lins, 1946) agree and feel that any tests of achievement in areas related to the future work of a teacher provide useful information concerning the general level of performance that may be expected in those areas in the future. Based on this, it is believed that external teachers examinations, intelligently interpreted enable the colleges to view their graduates, and the school their teachers, in relation to persons of comparable educational exposure throughout the country.

The intent of this study was to determine if there is any correlation between teacher effectiveness (based on criteria generally used to denote effectiveness) and NTE scores. The effectiveness was determined by the building principal's evaluation. This researcher is aware of the extreme limitations, and criticisms, of using principal ratings to measure teacher

effectiveness. Nevertheless, a recent state wide survey conducted in Alabama found (Brown & Maxson, 1972) that principal rating were used more than all other types of evaluation combined to assess teacher effectiveness.

If principal ratings are so widely used to measure the effectiveness of teachers it appears it would be beneficial to determine if there exists any relationship between their rating on selected teacher characteristics (which have been commonly listed as qualities of effective teachers) and NTE scores.

Quirk et al. (1973) stated in their review of the literature concerning the validity of the NTE that they had been unable to locate a single study that used scores earned on any of the Teaching Area Examinations (TAE) after 1964, when these scores were first equated to each other. The hypotheses tested in this investigation were: (1) to determine if there is a relationship between TAE scores and teacher effectiveness of secondary teachers as determined by the secondary building principal's evaluation; (2) to determine if there is a relationship between TAE scores and teacher effectiveness of elementary teachers as determined by the elementary building principal's evaluation; (3) to determine if there is any relationship between Weighter Common Examination Total (WCET) scores and selected characteristics of secondary teachers as determined by the secondary building principal's evaluation; and (4) to determine if there is any relationship between WCET scores

and selected characteristics of elementary teachers as determined by the elementary building principal's evaluation.

A review of the literature was conducted to determine an additional need for the study. Validation data are scarce for reasons that seem to be well understood by those who have tried to evaluate success in teaching. Crombach (1970) discussed a common type of criterion - the rating. He concluded that ratings are not entirely satisfactory as criteria. The rater may not know the facts about the person. Often a rating reflects the personal relation between man and supervisor rather than the quality of the man's work.

On the job ratings were studied by the Alabama Education Study Commission in the field of teaching. A memorandum was released on June 17, 1971 by the commission in which they stated that they were publishing extensive data reflecting the results of National Teacher Examinations taken by 11,500 graduating students of twenty-two teacher colleges in Alabama in 1968, 1969, and 1970.

This report was met with extensive criticism which was reflected in some twenty newspaper articles throughout Alabama dating from June 18, through August 21, 1971. The commission reported that results of the NTE from 1968 through 1970 showed an average rating of Alabama students of thrity on a scale with a top rating of one hundred and that the average rating of the twenty-two Alabama teacher colleges ranged from an average high of sixty-eight at the University of Alabama at Huntsville to a low of an average three at Alabama State University.

The President of Birmingham Southern, one of the schools scoring highest on the NTE, said that the scores indicate that teachers trained at Southern have a wide and well-rounded background. Dr. William Hunter, chairman of the School of Education at Tuskegee Institute and immediate past President of AACTE, said he has felt for some time that the NTE is slanted toward a particular educational philosophy, and that it lacks balance on points of view.

A determination was made by the investigator that a need existed to provide statistical data regarding the relationships between the TAE and the WCET, and teacher effectiveness as measured by principal evaluations. The four basic hypotheses, previously stated, were developed to test principal ratings and their relationships to the TAE and the WCET.

A rating form was constructed by the investigator and mailed to each selected teacher's principal (n = 66). A 90% return was experienced giving a final sample size of 59 (elementary = 22; secondary = 37). Each teacher's name had been obtained from the list of education graduates that had taken the NTE at one predominately white and one predominately black university in Alabama. Each subject used in this study had a minimum of three years teaching experience.

Analysis of the data was accomplished by determining relationship through the correlation technique. The significance level used for the acceptance-rejection decision regarding the stated hypotheses was set at .05. The Weighted Common Examinations total score from the NTE were used without any

transformations. The WCE has been described as a test "to measure general knowledge and ability requisite to effective teaching (Brownell, 1959)." Teaching Area Examination scores were also used without any transformations. The TAE "measures your understanding of subject matter and methods applicable to (various) subject areas . . . (NTE Bulletin, 1972-73, p.3)." Teachers were noted by their principals on a five point scale for 15 personal characteristics for effective teaching (motivation, classroom control, creativity, knowledge of subject matter, innovativeness, adaptiveness, professionalism, empathy, cooperativeness, morale, imaginativeness, tolerance, self-improvement, dependability, and overall effectiveness). The data regarding the principal's ratings and WCET scores and the principal's ratings and TAE scores were summed individually so as to produce an overall performance rating.

The means, standard deviations, and correlations were computed for the four variables. The results are presented in Tables 1 - 5.

Table 1

Principal Ratings of Teacher Effectiveness

Rater	M	SD
Elementary School Principal (n=)	4.4	1.98
Secondary School Principal (n=)	3.8	1.87

Table 2

WCET Scores (NTE)

Subjects	M	SD
Elementary Teachers	623.3	18.73
Secondary Teachers	643.4	19.02

Table 3

TAE Scores (NTE)

Subjects	M	SD
Elementary Teachers	556.5	18.17
Secondary Teachers	607.6	21.81

Table 4

Correlations - Principal Ratings and WCET Scores

Rater/Subject	df	r
Elementary Principals/Teachers	20	.38
Secondary Principals/Teachers	35	.41*

(*P $<$ .05)

Table 5

Correlations - Principal Ratings and TAE Scores

Rater/Subject	df	r
Elementary Principals/Teachers	20	.40
Secondary Principals/Teachers	35	.52**

(**P < .01)

The results indicate that the secondary principal may put more emphasis on knowledge of subject matter than does the elementary principal. For example, a secondary principal might say "he is a good algebra teacher", "where the elementary principal would say, "he is a good teacher", thus perhaps explaining the significant relationship found between principals ratings and TAE scores of secondary teachers and the lack of a significant relationship in the other areas.

It appears that a significant relationship does exist between TAE scores and secondary principal's evaluation, and between WCET scores and secondary principal's evaluation. However, no such significant relationship was found to exist concerning hypotheses two and four. A word of caution is in order. The NTE is a test of knowledge of the principles of teaching, not a test of the act of teaching. As yet there is no valid and reliable criterion for good teaching.

The last point I wish to make concerns subsequent research in this area. It occurred to this writer during the investigation that the crucial cognitive factor in teaching effectiveness is not intellectual aptitude or achievement as such but the teacher's general and special abilities as related to 1) the subject matter of the lesson, and 2) the level of each child in the class. For example, it may be found that a teacher who has average knowledge about the science of flowering plants has high effectiveness with the bright children in an elementary school science class, poor effectiveness with tne whole of a

junior high-school class in a culturally-deprived neighborhood, and, once again, high effectiveness with the slower children in a high-school class. If such a complex interaction exists between the subject matter and the abilities of the child and teacher, and it does seem plausible, then the teacher's effectiveness must be weighed relative to a group of children with a given ability in each subject matter area.

REFERENCES

Boozer, H. External exams as predictors of competence. Journal of Teacher Education, April, 1968, p. 214.

Brownell, W. Review of the National Teacher Examinations (in Bunas). The Fifth Mental Measurements Yearbook. Highland Park, N.J.: The Gnyphon Press, 1959.

Brown, S. and Maxson, R. Who evaluates teachers? (unpublished monograph) Auburn University at Montgomery, 1972.

Crombach, L. Essentials of psychological testing. New York: Harper and Row, 1970.

Cramer, F. Testing programs - misconceptions, misuse, overuse. Michigan Journal of Secondary Education, Spring, 1961, pp. 155-158.

Lins, L. The prediction of teaching efficiency. Journal of Experimental Education, April, 1946, pp. 2-60.

NTE Bulletin, 1972-73, p.3.

Quirk, T., Whitten, B., and Weinberg, S. A critical review of research related to the National Teacher Examinations. Princeton: Educational Testing Service, 1972.

Seashore, H. National Teacher Examinations (in Boston). The Sixth Mental Measurements Yearbook. Highland Park, N.J.: Gnyphon Press, 1965.